

Each Data Services server can contribute one — and only one — Job Server to a specific Server Group. Each Job Server collects resource utilization information for its computer. Data Services uses this information to determine where to execute a job, data flow, or subdata flow (depending on the distribution level specified).

The Data Services Engines

When Data Services jobs are executed, the Job Server starts Data Services engine processes to perform data extraction, transformation, and movement. Data Services engine processes use parallel processing and in-memory data transformations to deliver high data throughput and scalability.

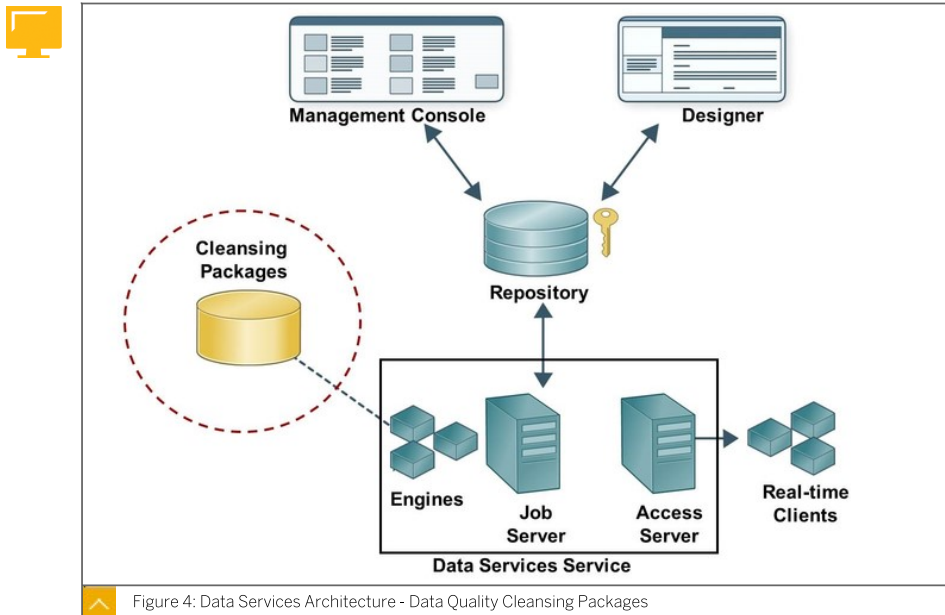


Figure 4: Data Services Architecture - Data Quality Cleansing Packages

The Data Quality Cleansing Packages files provide referential data for the Data Cleanse and Address Cleanse transforms to use when parsing, standardizing, and cleansing name and address data. Cleansing Packages enhance the ability of Data Cleanse to accurately process various forms of global data by including language-specific reference data and parsing rules.

They provide information on addresses from postal authorities; cleansing package files are used to identify, parse, and standardize data such as titles, firm data, acronyms, match standards, gender, capitalization, and address information.

The Data Services Management Console

The Data Services Management Console provides access to these features:

- Administration of Data Services resources, including:
 - Scheduling, monitoring, and executing batch jobs
 - Configuring, starting, and stopping real-time services